Prenatal Exposure to Maternal Cigarette Smoking and DNA Methylation: Epigenome-Wide Association in a Discovery Sample of Adolescents and Replication in an Independent Cohort at Birth through 17 Years of Age

Ken W. K. Lee, Rebecca Richmond, Pingzhao Hu, Leon French, Jean Shin, Celine Bourdon, Eva Reischl, Melanie Waldenberger, Sonja Zeilinger, Tom Gaunt, Wendy McArdle, Susan Ring, Geoff Woodward, Luigi Bouchard, Daniel Gaudet, George Davey-Smith, Caroline Relton, Tomas Paus, and Zdenka Pausova

This article will be available in its final, 508-conformant form 2–4 months after Advance Publication. If you require assistance accessing this article before then, please contact Dorothy L. Ritter, *EHP* Web Editor. *EHP* will provide an accessible version within 3 working days of request.

Prenatal Exposure to Maternal Cigarette Smoking and DNA Methylation: Epigenome-Wide Association in a Discovery Sample of Adolescents and Replication in an Independent Cohort at Birth through 17 Years of Age

Ken W. K. Lee,[1] Rebecca Richmond,[2] Pingzhao Hu,[1,3] Leon French,[4] Jean Shin,[1] Celine Bourdon,[1] Eva Reischl,[5] Melanie Waldenberger,[5] Sonja Zeilinger,[5] Tom Gaunt,[2] Wendy McArdle,[2] Susan Ring,[2] Geoff Woodward,[2] Luigi Bouchard,[6,7] Daniel Gaudet,[7,8] George Davey-Smith,[2] Caroline Relton,[9] Tomas Paus,[4] and Zdenka Pausova[1]

[1]The Hospital for Sick Children, University of Toronto, Toronto, Canada; [2]School of Social and Community Medicine, University of Bristol, Bristol, UK; [3]Department of Biochemistry and Medical Genetics, University of Manitoba, Winnipeg, Canada; [4]Rotman Research Institute, University of Toronto, Toronto, Canada; [5]Research Unit of Molecular Epidemiology, Helmholtz Zentrum Muenchen, Munich, Germany and Institute of Epidemiology II, Helmholtz Zentrum Muenchen, German Research Center for Environmental Health, Neuherberg, Germany; [6]Department of Biochemistry, Université de Sherbrooke, Sherbrooke, Canada; [7]ECOGENE-21 and Lipid Clinic, Chicoutimi Hospital, Chicoutimi, Canada; [8]Department of Medicine, Université de Montréal, Montréal, Canada; [9]Institute of Genetic Medicine, Newcastle University, Newcastle upon Tyne, UK

Address correspondence to Zdenka Pausova, The Hospital for Sick Children, Associate Professor, Departments of Physiology and Nutritional Sciences, University of Toronto, Toronto,

Canada. Telephone: (416) 813-7654/304340. Fax: (416) 813-5771. E-mail:

zdenka.pausova@sickkids.ca

Short title: Prenatal smoke exposure and DNA methylation

Competing financial interests: The authors declare they have no actual or potential competing financial interests.

Abstract

Background: Prenatal exposure to maternal cigarette smoking (prenatal smoke exposure) had been associated with altered DNA methylation (DNAm) at birth.

Objective: We examined whether such alterations are present from birth through adolescence.

Methods: We used the Infinium HumanMethylation450K BeadChip to search across 473,395 CpGs for differential DNAm associated with prenatal smoke exposure during adolescence in a discovery cohort (n=132) and at birth, during childhood, and during adolescence in a replication cohort (n=447).

Results: In the discovery cohort, we found five CpGs in *MYO1G* (top-ranking CpG: cg12803068, $p=3.3 \times 10^{-11}$) and *CNTNAP2* (cg25949550, $p=4.0 \times 10^{-9}$) to be differentially methylated between exposed and non-exposed individuals during adolescence. The CpGs in *MYO1G* and *CNTNAP2* were associated, respectively, with higher and lower DNAm in exposed vs. non-exposed adolescents. The same CpGs were differentially methylated at birth, during childhood, and during adolescence in the replication cohort. In both cohorts and at all developmental time-points, the differential DNAm was in the same direction and of a similar magnitude, and was not altered appreciably by adjustment for current smoking by the participants or their parents. In addition, four of the five EWAS-significant CpGs in the adolescent discovery cohort were also among the top sites of differential methylation in a previous birth cohort, and differential methylation of CpGs in *CYP1A1*, *AHRR* and *GFI1* observed in that study was also evident in our discovery cohort.

Conclusions: Our findings suggest that modifications of DNAm associated with prenatal maternal smoking may persist in exposed offspring for many years - at least until adolescence.

Introduction

Maternal cigarette smoking during pregnancy was and still is common in the industrialized world. In the US, for example, 30-40% of pregnant women smoked in the 1960s and 1970s and 16% of pregnant women smoke today (2012 US National Survey on Drug Use and Health). Prenatal exposure to maternal cigarette smoking (prenatal smoke exposure) has been associated with a number of health problems in the exposed offspring (Oken et al. 2008; Power et al. 2010; Syme et al. 2010). The underlying mechanisms of these associations are not clear but may involve lasting modulations of DNA methylation (DNAm).

DNAm is a chemical modification constituted most commonly by the addition of a methyl group to cytosines in CpG dinucleotides (CpGs) (Relton and Davey Smith 2010). It is catalyzed by two main types of DNA methyltransferases (DNMTs): *de novo* DNMTs, which play a key role in establishing new DNAm patterns during cell differentiation and embryogenesis, and maintenance DNMTs, which are responsible for copying these patterns from cell to cell during successive mitotic divisions (Jeltsch 2006; Tang et al. 2009). The main function of DNAm is the regulation of gene expression and genomic architecture (Levin and Moran 2011; Tsukahara et al. 2009).

Current research suggests that cigarette smoke is a powerful environmental modifier of DNAm (reviewed in (Lee and Pausova 2013)). Cigarette smoke contains a large number of chemicals, such as carcinogens, nicotine and carbon monoxide, that have been shown to modify DNAm in differentiating and dividing cells (Cuozzo et al. 2007; Di et al. 2012; Han et al. 2001; Huang et al. 2012; Mercer et al. 2009; Mortusewicz et al. 2005; Satta et al. 2008; Shahrzad et al. 2007). Consistent with these possible effects, reproducible differences in DNAm have been reported in

peripheral blood cells of smokers vs. non-smokers (Breitling et al. 2011; Philibert et al. 2012; Shenker et al. 2013; Zeilinger et al. 2013).

Many chemicals contained in cigarette smoke can easily pass from a smoking pregnant woman to the developing embryo/fetus (Lambers and Clark 1996). Accordingly, differences in DNAm have been reported in cord blood-cells of babies born to smoking vs. non-smoking mothers (Joubert et al. 2012). Whether these DNAm differences persist postnatally in peripheral blood cells of the exposed offspring has not been studied; this is the main subject of the present investigation. Prenatal exposure to cigarette smoke may induce *lasting* and *soma-wide (present in all somatic cells)* modifications of DNAm in the exposed offspring, particularly if occurring during early stages of embryogenesis when global erasure and re-establishment of DNAm patterns take place in yet-undifferentiated stem cells (Feng et al. 2010; Smith et al. 2012). These DNAm modifications are then propagated during embryonic development to all somatic cell lineages (including precursors of peripheral blood cells) and maintained throughout life by the action of maintenance DNMTs, which copy these modifications from cell to cell during successive cell divisions (Davies et al. 2012; Faulk and Dolinoy 2011; Jurkowska et al. 2011; Lee and Pausova 2013; Petronis 2010).

To investigate whether DNAm modifications induced by prenatal smoke exposure persist postnatally (at least until adolescence), we first conducted an epigenome-wide association study (EWAS) to search for CpGs methylated differentially between exposed and non-exposed adolescents in one cohort, the Saguenay Youth Study (SYS) (Pausova et al. 2007). Subsequently, we tested whether the identified CpGs were also methylated differentially between exposed and non-exposed newborns, children, and adolescents in another cohort, the Avon Longitudinal Study of Parents and Children (ALSPAC) (Boyd et al. 2013).

Methods

Discovery cohort: The Saguenay Youth Study (SYS)

Recruitment and assessment of prenatal smoke exposure

The SYS (n=1,028, 12- to 18-year old) is a population-based cross-sectional study of cardiovascular, metabolic, brain and mental health in adolescents carried out in the genetic founder population of the Saguenay Lac St. Jean region of Quebec, Canada (Pausova et al. 2007). All participants were White Caucasians. Adolescents exposed (n=490) and non-exposed (n=538) to maternal cigarette smoking during pregnancy were matched at recruitment by maternal education and school attended to reduce potential confounding related to socioeconomic status (Pausova et al. 2007). They were recruited via high schools. Exposure status was defined as having a mother who smoked at least 1 cigarette per day during the second trimester of pregnancy. Being non-exposed was defined as having a mother who did not smoke at least 1 year before and throughout the pregnancy. The information on exposure was obtained from the mother using a structured telephone interview at the time of recruitment of her adolescent offspring, and was validated subsequently against medical records from the time of pregnancy for a subset of adolescents; Kappa statistics with a mean (standard deviation) value of 0.69 (0.04) indicated a good strength of agreement in this subset (good agreement, 0.6-0.8) (Landis and Koch 1977). The main exclusion criteria were as follows: premature birth (<35 weeks) or detached placenta; maternal alcohol abuse during pregnancy (>210 ml of alcohol/week); positive medical history of type 1 diabetes, or of heart disease requiring surgery or sustained medication; and contraindications for magnetic resonance imaging. Additional recruitment details and selection criteria have been described previously (Pausova et al. 2007).

In the present study, we selected 66 exposed adolescents matched to 66 non-exposed by sex, age, and maternal education. The exposed adolescents were selected within each sex separately as those with the exposure being close to the sex-specific median (10 cigarettes/day in both sexes). This process resulted in a selection of 66 adolescents (33 males and 33 females) who were exposed to maternal cigarette smoking throughout pregnancy (all trimesters) at a level of 9±4 cigarettes/day (range 5-20 cigarettes/day). The studied adolescents were 12- to 18-year old (Table 1). Written consent of the parents and assent of the adolescents were obtained. The research ethics committees of the Chicoutimi Hospital and the Hospital for Sick Children approved the study.

Adolescent and parent cigarette smoking

Adolescents self-reported their mental health and substance use (including cigarette smoking) by completing a questionnaire developed for the SYS by J. Séguin from the University of Montreal (Groupe de Recherche sur l'Inadaptation Psychosociale); this questionnaire is based on validated National Longitudinal Survey of Children and Youth (NLSCY) and Quebec Longitudinal Study of Child Development (QLSCD) protocols (Pausova et al. 2007). Blood collection and adolescent smoking assessment occurred on the same day. Parents self-reported their own mental health and substance use (including cigarette smoking) using a similar questionnaire, as adapted by colleagues at the Groupe de Recherche sur l'Inadaptation Psychosociale of the University of Montreal (Pausova et al. 2007).

Epityping with the Infinium HumanMethylation450K BeadChip

DNA samples were extracted from the buffy coat (peripheral blood cells) using the Qiagen Autopure LS (Qiagen, Venlo, Netherlands). Bisulfite conversion was performed on 800 ng of

DNA from each sample with the EZ-96 DNA Methylation Kit (Zymo Research, Irvine, CA). DNA was subsequently used for hybridization on the Infinium HumanMethylation450K BeadChip (Illumina, San Diego, CA). Bisulfite conversion and methylation array hybridization were performed at the Helmholtz Zentrum München German Research Center for Environmental Health (Neuherberg, Germany). The Infinium HumanMethylation450K BeadChip interrogates methylation at >485,000 CpGs and provides coverage of > 99% of RefSeq genes (Sandoval et al. 2011). The DNAm score at each CpG, referred to as the DNAm β value, is derived from the fluorescent intensity ratio [β = intensity of the methylated allele / (intensity of the unmethylated allele + intensity of the methylated allele + 100)] (Bibikova et al. 2011). All samples were randomly loaded onto 11 arrays (12 samples per array) and processed by the same technician at the same time to minimize batch effects. Methylation values were normalized using the pre-process Illumina algorithm implemented in Minfi R package (Aryee et al. 2014). Parameters were set to mimic the Illumina Genome Studio normalization procedure (Illumina, San Diego, CA). Quality control was performed by excluding CpGs with detection $p \geq 0.05$ in more than 20% of samples (764 CpGs). After excluding these probes, as well as control probes and probes on sex chromosomes, a total of 473,395 CpGs were analyzed. All 132 samples had > 98% sites with detection $p < 0.05$.

Statistical methods (EWAS)

Methylation β values at each of the 473,395 CpGs were transformed to obtain M-values, defined as $\log_2 [\beta / (1 - \beta)]$, which have greater statistical power than β values to detect differential methylation at highly methylated and unmethylated CpGs (Du et al. 2010). Multivariable linear regression was used to perform association tests between the M value at each CpG (n=475,395) as the dependent variable and prenatal smoke exposure (exposed vs. non-exposed) as the

independent variable. These analyses were adjusted for potential confounding by age, sex, batch, and blood cell fractions (Model A). Blood cell fractions (granulocytes, B cells, CD8+ T cells, CD4+ T-cells, NK cells, monocytes) were determined by the method of Houseman (Houseman et al. 2012); the fractions were analyzed with principal component analysis and the first 3 components capturing 95% of variance were included in the regression model. The Manhattan and Q-Q plots for this Model A are shown in Supplemental Material, Figure S1. Note that none of the estimated blood cell fractions differed between exposed and non-exposed individuals (p=0.31-0.89). In additional analyses, we also adjusted for current smoking by adolescents (Model B) or for both current smoking by adolescents and second-hand smoking by parents (Model C). Further, we used the above-described basic Model A to analyze subsets of adolescents (a) who did not smoke themselves (61 exposed and 64 non-exposed, Model D) or (b) who did not smoke themselves and their parents did not smoke at present (19 exposed and 49 non-exposed, Model E). Finally, we ran an analysis that was not corrected for any potential confounders (Model F). CpGs achieving Bonferroni-corrected statistical significance of p<0.05 (uncorrected $p<1.1\times10^{-7}$) were considered EWAS significant. False discovery rate (FDR)-corrected p-values were also determined according to the method of Benjamini and Hochberg (Benjamini and Hochberg 1995). These statistical analyses were performed using the R software (http://www.r-project.org/).

Replication cohort: Avon Longitudinal Study of Parents and Children (ALSPAC)

Recruitment and assessment of prenatal smoke exposure

ALSPAC is a population-based longitudinal birth cohort ascertained in the former Avon Health Authority in southwest England between April 1st 1991 and December 31st 1992; the initial cohort included 14,541 pregnancies and 13,971 children alive at 12 months of age (Boyd et al.

2013). From this cohort, approximately 500 child participants have been epityped with the Infinium HumanMethylation450K BeadChip (Illumina, San Diego, CA) as part of a larger methylation project, ARIES, which is generating epigenetic information for 1,000 mother-offspring pairs at multiple time points across the life-course (http://www.ariesepigenomics.org.uk/). The individuals included in the present study are a subset of the offspring participants in ARIES for whom 450K data were available at the time of submission of this manuscript. Of these individuals, 82% (430/526) had methylation data available for at least two of the following 3 time points: at birth, 7 years of age, and 17 years of age. DNA samples were extracted from cord blood cells (birth) and peripheral blood cells (at 7 and 17 years of age). From these, we selected exposed individuals as child participants whose mothers smoked ≥ 5 cigarettes/day throughout their gestation and non-exposed individuals as child participants whose mothers did not smoke ~4/5 months before pregnancy and throughout all trimesters of their gestation (Supplemental Material, Table S1). Prenatal exposure to maternal cigarette smoking was assessed during pregnancy. Methylation values were normalized using the approach of Touleimat and Tost (Touleimat and Tost 2012).

Statistical methods

Associations between prenatal exposure to maternal cigarette smoking and M values at the 5 EWAS-significant CpGs identified in the SYS were tested using linear regression models adjusted for age, sex, maternal education, household social class, and batch. We included maternal education and household social class in the model, as unlike in the SYS exposed and non-exposed participants were not matched for any indices of socioeconomic status at recruitment. In additional analyses, we also adjusted for parent (second-hand) smoking at 7 years

of age, and for both participant smoking and parent (second-hand) smoking at 17 years of age. Statistical analyses were performed using the R software (http://www.r-project.org/).

Results

EWAS search for differential DNAm associated with prenatal smoke exposure in the SYS adolescents

The EWAS search was carried out in the SYS, which is a population-based cross-sectional study of 12- to 18-year old adolescents aimed at investigating long-term health outcomes associated with prenatal exposure to maternal cigarette smoking (Table 1). All participants are from the genetic founder population of the Saguenay Lac St. Jean region of Quebec, Canada (Pausova et al. 2007).

The EWAS search in the SYS identified 5 EWAS-significant ($p < 1.1 \times 10^{-7}$) CpGs (Figure 1A and Table 2, Model A). Four of them, including the most significant one (cg12803068, $p = 3.3 \times 10^{-11}$), were located within a 1-kb segment of the myosin IG gene (*MYO1G*); they were all associated with higher DNAm in exposed than non-exposed adolescents (Figure 1A and Table 2, Model A). DNAm at all 4 CpGs was correlated (r=0.83-0.92, p<0.0001) and as such the 4 CpGs are not independent signals. The other significant CpG was found in the contactin associated protein-like 2 gene (*CNTNAP2*, cg25949550, $p = 4.0 \times 10^{-9}$). In contrast with the *MYO1G* CpGs, it was associated with lower DNAm in exposed than non-exposed adolescents (Figure 1A and Table 2, Model A). None of the above CpGs were cross-reactive or polymorphic as listed in (Chen et al. 2013; Price et al. 2013). These results were obtained while adjusting for potential confounders, namely age, sex, batch, and blood cell fractions (Model A). They remained virtually unchanged when not adjusted for any of these potential confounders (Model F,

11

Supplemental Material, Table S2). In both adjusted and unadjusted analyses, differential DNAm (DNAm β difference between exposed and non-exposed) ranged from +0.04 to +0.09 in *MYO1G* and it was -0.02 in *CNTNAP2*.

Next, we examined whether current cigarette smoking by the adolescent participants or by their parents confounded these associations. In our sample, only 5 of 66 exposed and 2 of 66 non-exposed (Table 1) adolescents reported "smoking at least 1 cigarette in the last 30 days". Associations of prenatal maternal smoking and differential DNAm were comparable in models that also adjusted for participant smoking (Figure 1A, Model B) or that excluded current participant smokers (Figure 1A, Model D)." Adjusting for current smoking by parents as well as by the participants had little influence on point estimates, though p-values were increased somewhat: 2 of the 5 CpGs remained EWAS-significant, while the remaining 3 showed $p < 1.5 \times 10^{-6}$ (Figure 1A and Table 2, Model C). Excluding adolescents who were current smokers or whose parents were current smokers reduced the sample size from 66 exposed and 66 non-exposed to 18 exposed and 47 non-exposed. Nevertheless, all 5 CpGs remained significant (at $p < 0.05$) even in this smaller subsample (Figure 1A and Table 2, Model E).

Replication and persistence of differential DNAm associated with prenatal smoke exposure in the ALSPAC newborns, children and adolescents

The replication and persistence studies were carried out an independent sample, the ALSPAC cohort, which is a population-based longitudinal birth cohort ascertained in the former Avon Health Authority in southwest England between April 1st 1991 and December 31st 1992 (Supplemental Material, Table S1) (Boyd et al. 2013). Specifically, we tested whether the 5 EWAS-significant CpGs identified in the discovery SYS sample of adolescents, were also

differentially methylated in exposed vs. non-exposed newborns (39 vs. 402), children (7 years; 39 vs. 408), and adolescents (17 years, 13 vs. 206) from the ALSPAC cohort. We found that all 5 CpGs were associated with prenatal smoke exposure at birth and during childhood (Table 3, Models A' and B', respectively) – and 4 out of the 5 CpGs were associated with prenatal smoke exposure during adolescence (Table 3 Model D'). The direction and magnitude of differential DNAm at all 3 developmental time points for all 5 CpGs (Figure 1B) was consistent with the findings in the discovery cohort (Figure 1A). These results remained significant ($p < 0.05$) after adjusting for exposure to second-hand smoke (i.e., current smoking by parents) during childhood (Table 3, Model C') and for current participant smoking and parent smoking during adolescence (Table 3, Model E'), despite substantially reduced sample sizes (Figure 1B).

Discussion

The results of the present study suggest that prenatal smoke exposure is associated with modifications of DNAm that persist in the exposed offspring for years – from birth until at least adolescence – and that these associations remain after adjusting for blood-cell fractions, current smoking by adolescent participant, and current smoking by parents.

To our knowledge, our study is the first to report that prenatal smoke exposure-associated differential DNAm may persist in the exposed offspring for years after their birth. The MoBa study reported that prenatal smoke exposure was associated with differential DNAm at birth (Joubert et al. 2012). Importantly, all 5 CpGs we found to be differentially methylated in the present study at birth and during childhood (in the ALSPAC replication cohort), and during adolescence (in the SYS discovery sample and in the ALSPAC cohort) were also differentially methylated in the MoBa study at birth (Joubert et al. 2012) with the differences between exposed

and non-exposed being of the same direction and similar magnitude (Table 4). Other CpGs with prenatal smoke exposure-associated differential DNAm in the MoBa study were also significant at the strict look-up level (26 CpGs, $p < 1.92*10^{-3}$) in the SYS adolescents, including CpGs in *CYP1A1*, *AHRR*, and *GFI1*, and the gene-unmapped CpG 04598670 (Table 4). Differential DNAm across the 26 MoBa CpGs was correlated between the SYS and MoBa samples ($r^2 = 0.53$, $p < 0.0001$) (Supplemental Material, Figure S2)." Taken together, these studies indicate that prenatal smoke exposure is associated with specific DNAm modifications observed at birth (MoBa, ALSPAC at birth) and that some of these modifications may persist in the exposed offspring until at least adolescence (SYS, ALSPAC).

The EWAS-significant CpGs were located in *MYO1G* and *CNTNAP2*. *MYO1G* encodes a plasma membrane-associated class I myosin that is expressed abundantly in hematopoietic cells; it regulates cell elasticity and migration (Olety et al. 2010; Patino-Lopez et al. 2010). *CNTNAP2* encodes a member of the neurexin family of cell-adhesion molecules that plays a critical role in brain development (Anderson et al. 2012), and has been implicated in a number of neurodevelopmental disorders, including autism (Anney et al. 2012). One of the main functions of DNAm is the regulation of gene expression. The 4 EWAS-significant CpGs in *MYO1G* (but not the one in *CNTNAP2*) were located in a region containing binding sites for a number of transcription factors and enhancers (http://genome.ucsc.edu/ENCODE) and thus might alter mRNA expression of the gene. This possibility remains to be studied, however.

In addition to prenatal exposure to cigarette smoke, active smoking has been associated with differential DNAm in peripheral blood cells (Breitling et al. 2011; Philibert et al. 2012; Shenker et al. 2013; Zeilinger et al. 2013). In our discovery sample, only 5% of adolescents reported current cigarette smoking, and excluding these participants or adjusting for their current smoking

had little influence on model estimates. Further, we considered that postnatal *exposure to second-hand smoke* might be a confounder of the associations between prenatal smoke exposure and DNAm during childhood and adolescence. Although second-hand smoke exposure has not been reported to impact DNAm in peripheral blood cells, some evidence exists it is associated with differential DNAm in cancer tissues (though to a lesser degree than variations observed in active smokers) (Scesnaite et al. 2012; Wilhelm-Benartzi et al. 2011). Our results showed, however, that the differential DNAm associated with prenatal smoke exposure seen in the SYS adolescents was present (albeit somewhat attenuated) even after additional adjusting for second-hand smoke exposure to parental smoking. Importantly, the prenatal smoke exposure-associated differential DNAm we observed in the ALSPAC replication sample also remained significant after additional adjusting for adolescents' smoking and their exposure to second-hand smoke generated by the parents. Across both SYS (adolescence) and ALSPAC (birth, childhood, adolescence) samples and all statistical models we tested (with varying numbers of available participants), the magnitude of DNAm differences between exposed and non-exposed individuals was generally consistent. Collectively, these results support the postnatal persistence of prenatal smoke exposure-associated DNAm modifications in the exposed offspring.

In conclusion, our results suggest that prenatal smoke exposure may induce reproducible alterations of DNAm in the offspring, and that some of these alterations may persist for many years – well into adolescence. The potential implications of such epigenetic modifications for the "programming" of health require further study. Likewise, their possible utility as blood biomarkers of prenatal environmental exposures, such as maternal cigarette smoking, needs further exploration. In closing, promotion of smoking cessation remains imperative (Ng et al.

2014); gaining knowledge about molecular mechanisms capable of influencing health trajectories in the exposed population is critical for the prevention and treatment of associated disorders.

References

Anderson GR, Galfin T, Xu W, Aoto J, Malenka RC, Sudhof TC. 2012. Candidate autism gene screen identifies critical role for cell-adhesion molecule CASPR2 in dendritic arborization and spine development. Proc Natl Acad Sci U S A 109:18120-18125.

Anney R, Klei L, Pinto D, Almeida J, Bacchelli E, Baird G, et al. 2012. Individual common variants exert weak effects on the risk for autism spectrum disorders. Hum Mol Genet 21:4781-4792.

Aryee MJ, Jaffe AE, Corrada-Bravo H, Ladd-Acosta C, Feinberg AP, Hansen KD, et al. 2014. Minfi: A flexible and comprehensive bioconductor package for the analysis of Infinium DNA Methylation microarrays. Bioinformatics 30:1363-1369.

Benjamini Y, Hochberg Y. 1995. Controlling the false discovery rate: A practical and powerful approach to multiple testing. Journal of the Royal Statistical Society B 85:289-300.

Bibikova M, Barnes B, Tsan C, Ho V, Klotzle B, Le JM, et al. 2011. High density DNA methylation array with single CpG site resolution. Genomics 98:288-295.

Boyd A, Golding J, Macleod J, Lawlor DA, Fraser A, Henderson J, et al. 2013. Cohort profile: The 'children of the 90s'--the index offspring of the Avon Longitudinal Study of Parents and Children. Int J Epidemiol 42:111-127.

Breitling LP, Yang R, Korn B, Burwinkel B, Brenner H. 2011. Tobacco-smoking-related differential DNA methylation: 27k discovery and replication. Am J Hum Genet 88:450-457.

Chen YA, Lemire M, Choufani S, Butcher DT, Grafodatskaya D, Zanke BW, et al. 2013. Discovery of cross-reactive probes and polymorphic cpgs in the Illumina Infinium Human Methylation450 microarray. Epigenetics 8:203-209.

Cuozzo C, Porcellini A, Angrisano T, Morano A, Lee B, Di Pardo A, et al. 2007. DNA damage, homology-directed repair, and DNA methylation. PLoS Genet 3:e110.

Davies MN, Volta M, Pidsley R, Lunnon K, Dixit A, Lovestone S, et al. 2012. Functional annotation of the human brain methylome identifies tissue-specific epigenetic variation across brain and blood. Genome Biol 13:R43.

Di YP, Zhao J, Harper R. 2012. Cigarette smoke induces MUC5AC protein expression through the activation of SP1. J Biol Chem 287:27948-27958.

Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L, et al. 2010. Comparison of beta-value and M-value methods for quantifying methylation levels by microarray analysis. BMC Bioinformatics 11:587.

Faulk C, Dolinoy DC. 2011. Timing is everything: The when and how of environmentally induced changes in the epigenome of animals. Epigenetics 6:791-797.

Feng S, Jacobsen SE, Reik W. 2010. Epigenetic reprogramming in plant and animal development. Science 330:622-627.

Han L, Lin IG, Hsieh CL. 2001. Protein binding protects sites on stable episomes and in the chromosome from *de novo* methylation. Mol Cell Biol 21:3416-3424.

Houseman EA, Accomando WP, Koestler DC, Christensen BC, Marsit CJ, Nelson HH, et al. 2012. DNA methylation arrays as surrogate measures of cell mixture distribution. BMC Bioinformatics 13:86.

Huang J, Okuka M, Lu W, Tsibris JC, McLean MP, Keefe DL, et al. 2012. Telomere shortening and DNA damage of embryonic stem cells induced by cigarette smoke. Reprod Toxicol.

Jeltsch A. 2006. Molecular enzymology of mammalian DNA methyltransferases. Curr Top Microbiol Immunol 301:203-225.

Joubert BR, Haberg SE, Nilsen RM, Wang X, Vollset SE, Murphy SK, et al. 2012. 450k epigenome-wide scan identifies differential DNA methylation in newborns related to maternal smoking during pregnancy. Environ Health Perspect 120:1425-1431.

Jurkowska RZ, Jurkowski TP, Jeltsch A. 2011. Structure and function of mammalian DNA methyltransferases. Chembiochem 12:206-222.

Lambers DS, Clark KE. 1996. The maternal and fetal physiologic effects of nicotine. Semin Perinatol 20:115-126.

Landis JR, Koch GG. 1977. The measurement of observer agreement for categorical data. Biometrics 33:159-174.

Lee KW, Pausova Z. 2013. Cigarette smoking and DNA methylation. Front Genet 4:132.

Levin HL, Moran JV. 2011. Dynamic interactions between transposable elements and their hosts. Nat Rev Genet 12:615-627.

Mercer BA, Wallace AM, Brinckerhoff CE, D'Armiento JM. 2009. Identification of a cigarette smoke-responsive region in the distal MMP-1 promoter. Am J Respir Cell Mol Biol 40:4-12.

Mortusewicz O, Schermelleh L, Walter J, Cardoso MC, Leonhardt H. 2005. Recruitment of DNA methyltransferase into DNA repair sites. Proc Natl Acad Sci U S A 102:8905-8909.

Ng M, Freeman MK, Fleming TD, Robinson M, Dwyer-Lindgren L, Thomson B, et al. 2014. Smoking prevalence and cigarette consumption in 187 countries, 1980-2012. JAMA 311:183-192.

Oken E, Levitan EB, Gillman MW. 2008. Maternal smoking during pregnancy and child overweight: Systematic review and meta-analysis. Int J Obes 32:201-210.

Olety B, Walte M, Honnert U, Schillers H, Bahler M. 2010. Myosin 1G (MYO1G) is a haematopoietic specific myosin that localises to the plasma membrane and regulates cell elasticity. FEBS Lett 584:493-499.

Patino-Lopez G, Aravind L, Dong X, Kruhlak MJ, Ostap EM, Shaw S. 2010. Myosin 1G is an abundant class I myosin in lymphocytes whose localization at the plasma membrane depends on its ancient divergent pleckstrin homology (PH) domain (Myo1PH). J Biol Chem 285:8675-8686.

Pausova Z, Paus T, Abrahamowicz M, Almerigi J, Arbour N, Bernard M, et al. 2007. Genes, maternal smoking, and the offspring brain and body during adolescence: Design of the Saguenay Youth Study. Hum Brain Mapp 28:502-518.

Petronis A. 2010. Epigenetics as a unifying principle in the aetiology of complex traits and diseases. Nature 465:721-727.

Philibert RA, Beach SR, Brody GH. 2012. Demethylation of the aryl hydrocarbon receptor repressor as a biomarker for nascent smokers. Epigenetics 7:1331-1338.

Power C, Atherton K, Thomas C. 2010. Maternal smoking in pregnancy, adult adiposity and other risk factors for cardiovascular disease. Atherosclerosis 211:643-648.

Price ME, Cotton AM, Lam LL, Farre P, Emberly E, Brown CJ, et al. 2013. Additional annotation enhances potential for biologically-relevant analysis of the Illumina Infinium Humanmethylation450 Beadchip array. Epigenetics Chromatin 6:4.

Relton CL, Davey Smith G. 2010. Epigenetic epidemiology of common complex disease: Prospects for prediction, prevention, and treatment. PLoS Med 7:e1000356.

Sandoval J, Heyn H, Moran S, Serra-Musach J, Pujana MA, Bibikova M, et al. 2011. Validation of a DNA methylation microarray for 450,000 CpG sites in the human genome. Epigenetics 6:692-702.

Satta R, Maloku E, Zhubi A, Pibiri F, Hajos M, Costa E, et al. 2008. Nicotine decreases DNA methyltransferase 1 expression and glutamic acid decarboxylase 67 promoter methylation in GABAergic interneurons. Proc Natl Acad Sci U S A 105:16356-16361.

Scesnaite A, Jarmalaite S, Mutanen P, Anttila S, Nyberg F, Benhamou S, et al. 2012. Similar DNA methylation pattern in lung tumours from smokers and never-smokers with second-hand tobacco smoke exposure. Mutagenesis 27:423-429.

Shahrzad S, Bertrand K, Minhas K, Coomber BL. 2007. Induction of DNA hypomethylation by tumor hypoxia. Epigenetics 2:119-125.

Shenker NS, Polidoro S, van Veldhoven K, Sacerdote C, Ricceri F, Birrell MA, et al. 2013. Epigenome-wide association study in the European Prospective Investigation into Cancer and Nutrition (EPIC-Turin) identifies novel genetic loci associated with smoking. Hum Mol Genet 22:843-851.

Smith ZD, Chan MM, Mikkelsen TS, Gu H, Gnirke A, Regev A, et al. 2012. A unique regulatory phase of DNA methylation in the early mammalian embryo. Nature 484:339-344.

Syme C, Abrahamowicz M, Mahboubi A, Leonard GT, Perron M, Richer L, et al. 2010. Prenatal exposure to maternal cigarette smoking and accumulation of intra-abdominal fat during adolescence. Obesity (Silver Spring) 18:1021-1025.

Tang M, Xu W, Wang Q, Xiao W, Xu R. 2009. Potential of DNMT and its epigenetic regulation for lung cancer therapy. Curr Genomics 10:336-352.

Touleimat N, Tost J. 2012. Complete pipeline for Infinium Human Methylation 450k Beadchip data processing using subset quantile normalization for accurate DNA methylation estimation. Epigenomics 4:325-341.

Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T. 2009. Bursts of retrotransposition reproduced in arabidopsis. Nature 461:423-426.

Wilhelm-Benartzi CS, Christensen BC, Koestler DC, Houseman EA, Schned AR, Karagas MR, et al. 2011. Association of secondhand smoke exposures with DNA methylation in bladder carcinomas. Cancer Causes Control 22:1205-1213.

Zeilinger S, Kuhnel B, Klopp N, Baurecht H, Kleinschmidt A, Gieger C, et al. 2013. Tobacco smoking leads to extensive genome-wide changes in DNA methylation. PLoS One 8:e63812.

Figure Legend

Figure 1. EWAS-significant CpGs in the discovery SYS and replication ALSPAC cohorts.
Mean differences in DNAm β values (and their 95% confidence intervals) between exposed (E)
and non-exposed (NE) individuals during adolescence in the discovery SYS cohort and at birth,
during childhood and during adolescence in the replication ALSPAC cohort are presented. In all
presented models, the differences in the SYS cohort were adjusted for age, sex, batch and blood
cell fractions, and the differences in the ALSPAC cohort were adjusted for age, sex, batch,
maternal education and parental social class. SYS: Saguenay Youth Study; ALSPAC: Avon
Longitudinal Study of Parents and Children. P-values of the associations are presented in Tables
2 and 3.

Table 1. Basic characteristics of SYS adolescents (discovery sample).

Characteristic	Non-exposed	Exposed	P-value
Sex (males/females)	33/33	33/33	N/A*
Age (years)	15.6 ± 1.5	15.5 ± 1.4	0.95
Maternal education**	5.4 ± 1.9	5.1 ± 2.0	0.47
Family income ($CAD/year)	60,384 ± 24,387	58,939 ± 25,548	0.74
Current adolescent smoking (yes/no)#	2/61	5/60	0.44
Current paternal smoking (yes/no)##	11/52	19/43	0.08
Current maternal smoking (yes/no)†	5/61	39/27	<0.0001
Maternal smoking during pregnancy (cigarettes/day)	0	8.9 ± 3.9	N/A*
Maternal age at delivery (years)	27.0 ± 4.5	25.0 ± 4.3	0.01
Length of gestation (weeks)	39.6 ± 1.5	39.6 ± 1.6	0.83
Birth weight (g)	3,442 ± 406	3,275 ± 501	0.04

Data are presented as mean ± standard deviation.

* Determined by the selection of the EWAS-studied participants

** Maternal education was classified as the highest achieved level : (1) primary school not completed, (2) primary school completed, (3) high school not completed, (4) high school completed, (5) college completed, or (6) university completed.

#Current adolescent smoking was defined as smoking of at least 1 cigarette in the last 30 days (data missing in 4 adolescents).

Current paternal smoking was defined as smoking of at least 1 cigarette in the last 30 days (data missing in 7 fathers).

† Current maternal smoking was defined as smoking of at least 1 cigarette in the last 30 days.

Table 2. EWAS: prenatal smoke exposure-associated differential DNA methylation during adolescence in the discovery SYS sample.

CpG ID	Gene	Model A		Model B		Model C		Model D		Model E	
		DNAm β difference (SE)	P-value	DNAm β difference (SE)	P-value	DNAm β difference (SE)	P-value	DNAm β difference (SE)	P-value	DNAm β difference (SE)	P-value
cg12803068	MYO1G	0.09 (0.01)	3.3E-11*	0.09 (0.01)	7.0E-11*	0.10 (0.01)	2.2E-08*	0.09 (0.01)	4.9E-10*	0.09 (0.02)	2.2E-05
cg04180046	MYO1G	0.07 (0.01)	3.9E-11*	0.07 (0.01)	1.2E-10*	0.08 (0.01)	1.2E-08*	0.07 (0.01)	4.4E-10*	0.07 (0.02)	4.0E-05
cg19089201	MYO1G	0.04 (0.01)	1.6E-09*	0.04 (0.01)	4.0E-09*	0.05 (0.01)	3.5E-07	0.04 (0.01)	1.8E-08*	0.04 (0.01)	7.9E-05
cg25949550	CNTNAP2	-0.02 (0.00)	4.0E-09*	-0.02 (0.00)	8.2E-10*	-0.02 (0.00)	9.4E-07	-0.02 (0.00)	1.2E-08*	-0.01 (0.01)	1.6E-02
cg22132788	MYO1G	0.06 (0.01)	8.4E-09*	0.06 (0.01)	1.5E-08*	0.06 (0.01)	1.5E-06	0.05 (0.01)	4.6E-08*	0.04 (0.01)	1.3E-03
cg23222488	NA	0.03 (0.01)	1.2E-06	0.03 (0.01)	3.9E-06	0.03 (0.01)	3.7E-05	0.02 (0.01)	1.2E-05	0.01 (0.01)	8.6E-03
cg27214774	FAM65C	0.01 (0.00)	1.9E-06	0.01 (0.00)	3.3E-06	0.01 (0.00)	9.3E-05	0.01 (0.00)	1.9E-05	0.01 (0.00)	7.6E-02
cg00253658	NA	0.07 (0.01)	2.0E-06	0.07 (0.01)	3.3E-06	0.07 (0.01)	8.8E-05	0.07 (0.01)	5.9E-06	0.05 (0.03)	1.3E-03
cg27641820	ANKRD11	-0.01 (0.00)	5.3E-06	0.00 (0.00)	8.2E-06	-0.00 (0.00)	2.1E-04	-0.01 (0.00)	5.4E-06	-0.00 (0.00)	1.3E-01
cg13323902	VTRNA1-1	-0.01 (0.00)	6.6E-06	-0.01 (0.01)	3.8E-06	-0.01 (0.01)	2.4E-04	-0.02 (0.01)	7.9E-07	-0.03 (0.01)	1.2E-05
cg05176844	HCG27	-0.02 (0.00)	7.2E-06	-0.02 (0.00)	4.1E-06	-0.02 (0.00)	1.3E-05	-0.02 (0.01)	2.1E-06	-0.02 (0.01)	1.0E-03
cg05549655	CYP1A1	0.03 (0.01)	7.6E-06	0.03 (0.01)	4.5E-06	0.03 (0.01)	4.6E-04	0.03 (0.01)	6.5E-06	0.02 (0.02)	1.9E-02
cg01586011	NA	0.02 (0.01)	8.3E-06	0.01 (0.01)	4.0E-05	0.01 (0.01)	6.3E-03	0.01 (0.01)	1.0E-04	-0.01 (0.01)	2.1E-01
cg17293195	KRTAP5-7	-0.02 (0.01)	1.2E-05	-0.02 (0.01)	8.8E-06	-0.02 (0.01)	1.0E-03	-0.02 (0.01)	4.2E-05	-0.02 (0.01)	4.5E-02
cg17924476	AHRR	0.03 (0.01)	1.2E-05	0.03 (0.01)	5.0E-05	0.03 (0.01)	6.6E-04	0.03 (0.01)	1.8E-04	0.03 (0.01)	1.2E-01
cg13570656	CYP1A1	0.04 (0.01)	1.3E-05	0.04 (0.01)	5.3E-06	0.04 (0.01)	2.9E-04	0.04 (0.01)	1.1E-05	0.00 (0.02)	1.2E-02
cg15982595	C13orf35	-0.01 (0.00)	1.5E-05	-0.01 (0.00)	3.0E-05	-0.01 (0.00)	1.7E-03	-0.01 (0.00)	3.3E-05	-0.01 (0.00)	6.1E-02
cg17101494	NA	-0.04 (0.02)	1.6E-05	-0.04 (0.02)	1.6E-05	-0.04 (0.02)	2.8E-04	-0.04 (0.02)	1.9E-05	-0.04 (0.02)	6.0E-04
cg11936410	TPM1	-0.01 (0.00)	1.7E-05	-0.01 (0.00)	9.2E-06	-0.01 (0.00)	1.5E-03	-0.01 (0.00)	1.7E-05	-0.01 (0.00)	3.1E-04
cg00213123	CYP1A1	0.02 (0.01)	1.9E-05	0.02 (0.01)	1.2E-05	0.02 (0.01)	8.1E-04	0.02 (0.01)	1.5E-05	0.00 (0.01)	2.3E-02
cg13418795	GUCY2D	-0.01 (0.00)	1.9E-05	-0.01 (0.00)	8.8E-05	-0.01 (0.00)	9.9E-05	-0.01 (0.00)	2.9E-05	-0.01 (0.01)	2.7E-03
cg05552796	NA	0.00 (0.00)	2.0E-05	0.00 (0.00)	2.4E-05	0.01 (0.00)	1.2E-04	0.00 (0.00)	1.4E-04	-0.00 (0.00)	1.9E-02
cg04515311	NA	-0.01 (0.00)	2.1E-05	-0.01 (0.00)	1.0E-04	-0.01 (0.00)	4.5E-03	-0.01 (0.00)	9.4E-05	-0.01 (0.01)	7.4E-02
cg12245321	NA	-0.01 (0.00)	2.1E-05	-0.01 (0.00)	2.8E-05	-0.01 (0.00)	2.1E-03	-0.01 (0.00)	3.3E-05	-0.01 (0.01)	9.3E-04
cg07719772	NA	0.02 (0.01)	2.7E-05	0.02 (0.01)	1.4E-05	0.02 (0.01)	1.3E-03	0.02 (0.01)	4.9E-06	0.01 (0.01)	1.1E-01

Model A: Adjusted for age, sex, batch and blood cell fractions (66 exposed vs. 66 non-exposed).

Model B: Adjusted for current participant smoking in addition to age, sex, batch and blood cell fractions (63 exposed vs. 65 non-exposed).

Model C: Adjusted for current participant and parent smoking in addition to age, sex, batch and blood cell fractions (62 exposed vs. 63 non-exposed).

Model D: Current participant smokers excluded; adjusted for age, sex, batch and blood cell fractions (60 exposed vs. 61 non-exposed).

Model E: Current participant and/or parent smokers excluded; adjusted for age, sex, batch and blood cell fractions (18 exposed vs. 47 non-exposed).

Differential DNA methylation (DNAm) between adolescents exposed and non-exposed prenatally to maternal cigarette smoking was assessed with linear regression using M-values at 473,395 CpGs. The results are presented as adjusted mean differences in DNAm β values between exposed and non-exposed individuals (DNAm β difference) and their respective standard errors (SEs) and uncorrected (nominal) P-values from the linear regression described above. *Indicates FDR-significance ($p<0.05$). Chromosome position (indicated by chromosome number followed by position) is based on the NCBI Human Reference Genome Assembly Build 37.p10.

Table 3. Prenatal smoke exposure-associated differential DNA methylation (DNAm) at EWAS-significant CpGs from the discovery SYS sample in newborns, children and adolescents in the replication ALSPAC sample.

CpG ID	Gene	Model A'		Model B'		Model C'		Model D'		Model E'	
		DNAm β difference (SE)	P-value	DNAm β difference (SE)	P-value	DNAm β difference (SE)	P-value	DNAm β difference (SE)	P-value	DNAm β difference (SE)	P-value
cg12803068	*MYO1G*	0.15 (0.03)	5.4E-10	0.17 (0.03)	4.8E-11	0.18 (0.04)	1.6E-07	0.19 (0.05)	1.7E-05	0.18 (0.07)	3.4E-03
cg04180046	*MYO1G*	0.08 (0.01)	7.7E-12	0.04 (0.01)	2.1E-08	0.05 (0.01)	4.9E-07	0.06 (0.02)	6.8E-04	0.07 (0.03)	1.1E-02
cg19089201	*MYO1G*	0.03 (0.01)	2.6E-04	0.06 (0.01)	2.2E-07	0.05 (0.02)	6.8E-04	0.07 (0.03)	3.8E-03	0.10 (0.04)	1.2E-02
cg25949550	*CNTNAP2*	-0.01 (0.00)	6.3E-07	-0.01 (0.00)	4.0E-04	-0.01 (0.00)	2.6E-03	-0.00 (0.00)	1.6E-01	0.00 (0.01)	8.7E-01
cg22132788	*MYO1G*	0.05 (0.01)	4.0E-10	0.09 (0.02)	5.3E-13	0.09 (0.02)	3.9E-08	0.06 (0.03)	1.9E-03	0.08 0.04	2.8E-03

Model A': At birth (39 exposed vs.402 non-exposed).

Model B': At age 7 years (39 exposed vs. 408 non-exposed).

Model C': At age 7 years (36 exposed vs. 380 non-exposed); additionally adjusted for current parent smoking.

Model D': At age 17 years (13 exposed vs. 206 non-exposed).

Model E': At age 17 years (8 exposed vs. 108 non-exposed); additionally adjusted for current parent and adolescent smoking

Differential DNAm between individuals exposed and non-exposed prenatally to maternal cigarette smoking was assessed with linear regression using M-values at 5 CpGs that were EWAS-significant in the SYS discovery cohort. All models (A-E) were adjusted for age, sex, maternal education, household social class and batch. The results are presented as adjusted mean differences in DNAm β values between exposed and non-exposed individuals (DNAm β difference) and their respective standard errors (SEs) and uncorrected (nominal) P-values from the linear regression described above. Chromosome position (indicated by chromosome number followed by position) is based on the NCBI Human Reference Genome Assembly Build 37.p10.
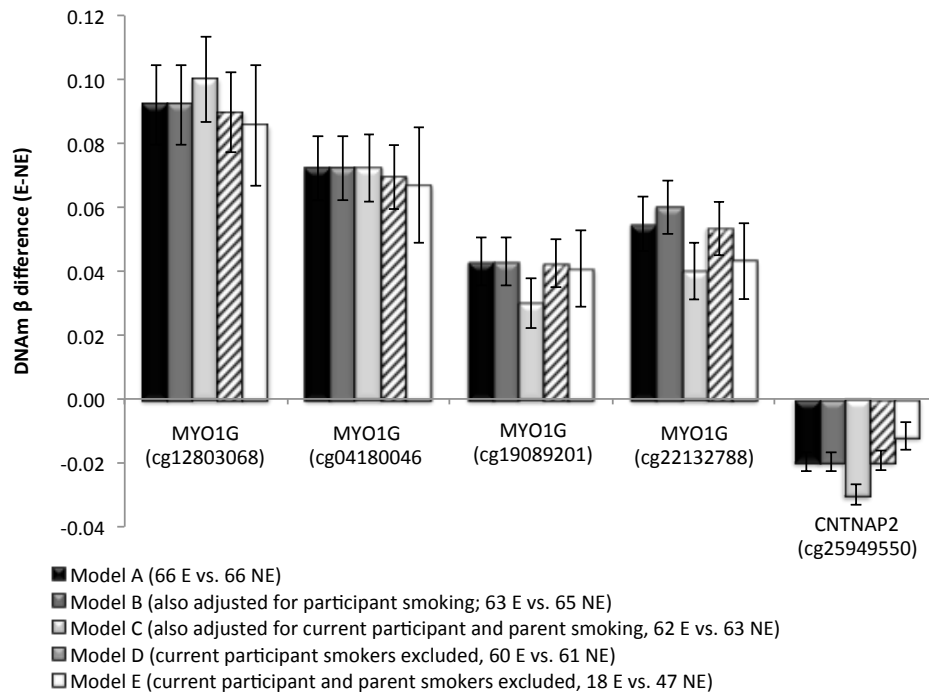
ALSPAC: Avon Longitudinal Study of Parents and Children

Table 4. Look-up in the SYS discovery sample of adolescents of the 26 CpGs identified previously at p<1x10[-7] in the MoBa sample of newborns.

CpG ID	Chr	Gene	MoBa newborns*		SYS adolescents	
			DNAm β difference (%)	P-value	DNAm β difference (%)	P-value
cg12803068	7	*MYO1G*	8.3	8.51E-19	9.2	**3.34E-11**
cg22549041	15	*CYP1A1*	7.2	4.52E-09	4.3	**9.70E-05**
cg18092474	15	*CYP1A1*	5.9	1.10E-08	2.9	**3.60E-04**
cg04180046	7	*MYO1G*	5.3	8.76E-20	7.2	**3.90E-11**
cg12477880	21	*RUNX1*	4.6	1.02E-09	6.2	3.46E-03
cg05549655	15	*CYP1A1*	3.5	2.96E-10	2.9	**7.57E-06**
cg11924019	15	*CYP1A1*	3.2	2.62E-08	2.8	**3.35E-05**
cg23067299	5	*AHRR*	3.2	4.21E-10	1.1	**5.74E-04**
cg22132788	7	*MYO1G*	2.8	1.98E-18	5.5	**8.41E-09**
cg19089201	7	*MYO1G*	1.4	3.22E-10	4.3	**1.58E-09**
cg03346806	8	*NA*	-1.5	3.08E-08	-0.3	3.44E-01
cg25949550	7	*CNTNAP2*	-1.8	4.15E-30	-2.0	**3.97E-09**
cg18655025	14	*TTC7B*	-1.8	2.07E-08	-0.7	1.32E-01
cg03991871	5	*AHRR*	-2.2	2.04E-11	-2.6	7.17E-02
cg21161138	5	*AHRR*	-2.3	1.52E-11	-1.1	6.17E-02
cg04598670	7	*NA*	-3	1.29E-11	-3.4	**1.27E-04**
cg10399789	1	*GFI1*	-3.7	4.08E-13	0.6	2.09E-01
cg06338710	1	*GFI1*	-5.8	1.34E-18	0.5	5.94E-01
cg09662411	1	*GFI1*	-6.6	2.26E-20	0.0	8.37E-01
cg18316974	1	*GFI1*	-7.1	6.43E-24	0.5	8.90E-01
cg05575921	5	*AHRR*	-7.5	2.85E-39	-1.7	1.68E-02
cg11715943	6	*HLA-DPB2*	-7.5	1.00E-08	-0.7	7.10E-01
cg14179389	1	*GFI1*	-8.6	5.38E-28	-4.7	**5.13E-05**
cg12876356	1	*GFI1*	-11.9	2.29E-30	-0.5	6.63E-01
cg18146737	1	*GFI1*	-12.3	2.42E-30	0.6	8.30E-01
cg09935388	1	*GFI1*	-13.7	1.05E-38	-1.5	1.31E-01

*Joubert, B.R. et al. (2012) *Environ Health Perspect*, 120, 1425-1431. DNA methylation (DNAm) β value differences between exposed and non-exposed individuals are shown. CpGs replicated in the discovery SYS sample at strict look-up level significance (26 tests, p<1.92*10[-3]) are indicated in bold. The SYS Model A (adjusted for age, sex, batch and blood cell fractions) is presented.

A) SYS discovery cohort

Model A (66 E vs. 66 NE)
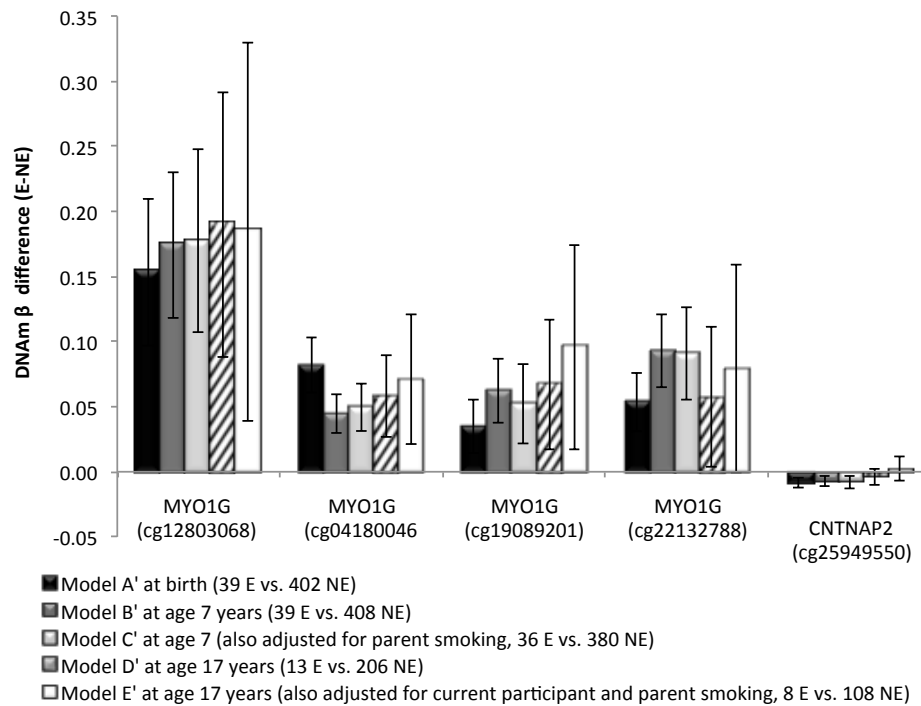Model B (also adjusted for participant smoking; 63 E vs. 65 NE)
Model C (also adjusted for current participant and parent smoking, 62 E vs. 63 NE)
Model D (current participant smokers excluded, 60 E vs. 61 NE)
Model E (current participant and parent smokers excluded, 18 E vs. 47 NE)

B) ALSPAC replication cohort

Model A' at birth (39 E vs. 402 NE)
Model B' at age 7 years (39 E vs. 408 NE)
Model C' at age 7 (also adjusted for parent smoking, 36 E vs. 380 NE)
Model D' at age 17 years (13 E vs. 206 NE)
Model E' at age 17 years (also adjusted for current participant and parent smoking, 8 E vs. 108 NE)

Figure 1.